

1108-62-38

Timothy D Rey* (trey1@steelcase.com), GH-4W, 901 44th Street, SE, Grand Rapids, MI 49508. *The Curse of Dimensionality in a Real Life Industrial Problem*. Preliminary report.

Industrial data mining (supervised learning) problems generally involve wrestling with the “curse of dimensionality”. Data collected in an industrial transaction environment is rarely if ever intended to be used in a modeling problem, let alone in a “cause and effect” modeling problem. Thus the contradiction between cause and effect (ala the use of the scientific method and proper design of experiments) and “prediction” is before us. This curse of dimensionality seems to prevent analytics professionals from finding true cause and effect. Data sets not intended for modeling generally have significant multicollinearity, lack of balance and often are too wide (p being inappropriate for n). Approaches for solving these issues can be broken down into three classes; dimension reduction, parameter adjustment and data structure adjustment. This talk will show an industrial data mining problem where the curse is present in all its glory. Each of the three basic methods for supposed “solutions” to the problem will be presented using modern day technologies. (Received December 02, 2014)